

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2001, 00:14:33 ; Search time 36.13 Seconds  
(without alignments)

1439.891 Million cell updates/sec

Title: US-09-494-297-2

Perfect score: 3945

Sequence: 1 MKKTRFPNKNTLNTQRYLS.....IAGISLGINGHTRIRKHD 757

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR.67:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	633.5	16.1	659	S40043	adhesin - Streptococcus
2	603	15.3	638	SS4418	fibronectin-binding
3	164	4.2	1039	T30856	protein F2 - Streptococcus
4	159	4.0	1243	S60138	sex factor aggregase
5	157.5	4.0	5005	F82884	hypothetical protein
6	151	3.8	2367	S70172	toxin B - Clostridium
7	149	3.8	1315	T28679	fibronectin-binding
8	148.5	3.8	1883	G82875	hypothetical protein
9	145	3.7	1036	T30311	S-layer protein - Streptococcus
10	144.5	3.7	1092	T30214	fibronectin-binding
11	143	3.6	1185	A42404	collagen adhesin - Streptococcus
12	142.5	3.6	627	T03821	prolipoprotein p65
13	139.5	3.5	888	E82885	hypothetical protein
14	139.5	3.5	1167	F71909	hypothetical protein
15	139	3.5	621	B64546	chaperone and heat shock
16	139	3.5	1433	A36734	toxin B - Clostridium
17	139	3.5	2366	S10317	surface layer protein
18	138.5	3.5	625	T03837	surface layer protein
19	138.5	3.5	2484	T26216	hypothetical protein
20	138.5	3.5	2607	T26215	hypothetical protein
21	137	3.5	1177	I64233	hypothetical protein
22	135	3.4	1117	S33851	fibronectin-binding
23	135	3.4	1176	A33856	surface-layer 125K
24	135	3.4	1228	I40468	surface layer protein
25	135	3.4	1301	S51323	SAC protein - Yersinia
26	135	3.4	1650	T18444	hypothetical protein
27	134	3.4	1668	A60272	Iga-specific metal
28	133	3.4	965	T32574	hypothetical protein
29	132.5	3.4	1116	A35129	surface layer protein

30	132.5	3.4	2287	T21312	hypothetical protein
31	132.5	3.4	3097	T00021	DN-cadherin - fruit
32	132	3.3	1103	H82884	multiple banded an
33	131.5	3.3	-867	D82940	conserved hypothet
34	131	3.3	843	S33442	EF protein - Streptococcus
35	131	3.3	1092	S42798	fibronectin-binding
36	131	3.3	1822	S33441	EF protein - Streptococcus
37	130.5	3.3	881	S37895	LH1 protein - Streptococcus
38	130.5	3.3	908	C70168	DNA polymerase I (
39	130.5	3.3	2166	G70163	hypothetical prote
40	130	3.3	621	A71961	90kDa chaperone -
41	130	3.3	1372	T43296	cell fusion protei
42	130	3.3	3381	T42389	versican precursor
43	130	3.3	5175	T20992	hypothetical prote
44	130	3.3	5198	T43290	hemiscientin precu
45	129.5	3.3	1684	S10789	amylase A-180 - al

## ALIGNMENTS

RESULT	1	
S40043		
adhesin - Streptococcus pyogenes		
C:Species: Streptococcus pyogenes		
C:Date: 19-Mar-1997 #sequence-revision 19-Mar-1997 #text-change 28-May-1999		
C:Accession: S40043		
R:Selal, S.; Avila, A.; Tovl, A.; Bursstein, I.; Caparon, G.; Hanski, E.		
Mol. Microbiol. 10, 1049-1055, 1993		
A:Title: Protein F: an adhesin of Streptococcus pyogenes binds fibronectin via two di		
A:Reference number: S40043; MUID:95020565		
A:Accession: S40043		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-659 <SEL>		
A:Cross-references: GB:LI0919; NID:9425479; PIDN:AAA26964.1; PID:9425480		
Query Match	16.1%;	Score 633.5; DB 2; Length 659;
Best Local Similarity	26.1%;	Pred. No. 6.1e-30;
Matches 203; Conservative 128; Mismatches 237; Indels 211; Gaps 30;		
QY 20 SKNSKREFTYLVGVFLMIRALVYSMGAKTVGCVSSSTPMAINPDSSSEYKRYGESVY 79		
DB 24 TKRRKRFVAVTLVGVFPMLLACAGAFGGVAAADKTVPSHSP--NPEFPVYGYDAIG 81		
QY 80 RGHPIYKQFRVAHDLRVNLGSRYSQVYCFNLKRAPIPLGSDSVKRWYKHHDSIFKFD 139		
DB 82 KEYPGYNITRTHDLRVNLGSRYSQVYCFNLKRAPIPLGSDSVKRWYKHHDSIFKFD 141		
QY 140 YANSPRTGDELNOKIRAVMYNGHPONANGIMEGLEPLNARIVTQDAVWYYSNADISNP 199		
DB 142 YATRTKLGKEELEQRIILSLYNAVYPDANGYMGLEHLNATVYQAVVHYSNSQYQF 200		
QY 200 DESFKRESNLSVTSQSLSMRQALKQIDPNATMPQVOPDPLSTFEESDKDKDXN 259		
DB 201 ETLWESAEAKGKISRSQVTLRREALKLLIDPNLEATVANKVIRGLNLFESL-----N 254		
QY 260 KGYONLISGLVTPKTPPGDPPMPNPOTTSVLIRKVAIGDYSKLLGATL----QLT 315		
DB 255 EAYONLISAEVYDPDPKPGGE--TSEHNKRTPE-----LQCTPIPEPKHP 298		
QY 316 GDNVNSFOARVSSNDIGRIELSDGYTLTEINSPAGYSIAEPTFKYKAGKRVYTIIDG 375		
DB 299 DDNLLEPTLPVMD---GEV-----PEVSSSLPALPLMPE-----LDG 337		
QY 376 KQLENKKEIVE-PYSEVAYNDFEESVLTQVYAFYAKKNNGSQVYCCNADLKSP 434		
DB 338 QEV--PEKPSIDPLIVPRE-----FNNKQOSP 364		
QY 435 PDEEDGKMTPTFTGEVY--THIAGRDLEKTYVPRPTDPTFLKHIKVLEKGYRKE 493		
DB 365 LAGE-----SGEYETITEVYGNQ-----QNPVIDK-----KLPRNTEGF--S 399		

[illegible]

RESULT 2

554418  
Fibronectin-binding protein precursor - Streptococcus pyogenes

C:Species: Streptococcus pyogenes  
C:Date: 06-Sep-1996 #sequence\_revision 13-Mar-1997 #text\_change 15-Oct-1999  
C:Accession: 554418; A44792; S28448

R:Title: S. R.; Valentin-Weigand, P.; Timmls, K.N.; Chatwal, G.S.  
MOL: Microbiol. 13, 531-539, 1994

A:Title: Domain structure and conserved epitopes of Sfb protein, the fibronectin-binding  
A:Reference number: 554418; MUID:95089690

A:Accession: 554418  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-638 <TAL>

A:Cross-references: EMBL:X67947; NID:G511149; PIDN:CAA48133.1; PID:G511150  
R:Title: S. R.; Valentin-Weigand, P.; Jerlstrom, P.G.; Timmls, K.N.; Chatwal, G.S.  
Infect. Immun. 60, 3837-3844, 1992

A:Title: Fibronectin-binding protein of Streptococcus pyogenes: sequence of the binding  
A:Reference number: A44792; MUID:92363585

A:Accession: A44792  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 371-638 <TA2>

A:Cross-references: EMBL:X67947  
A:Note: sequence extracted from NCBI backbone (NCBIN:111014, NCBIF:111015)

C:Genetics:  
A:Gene: sfb

C:Keywords: fibronectin binding  
F:1-49/Domain: signal predicted <SIG>  
F:50-638/Product: fibronectin-binding protein #status predicted <MAT>

[illegible]

QY	253	DKGKYNKXGONLLSGGVLPKPRTPEDDPMPMPNOPTVLLIKRYAIGYSGSLBEATL	312
Db	249	D-----KLIONLLSAEVPENPKPBE--TPREGRPTPEL-----DGTIIPBEP--	290
QY	313	QLTGDVNVNSFOARVSSNDIGERIELSDGTYTTLTSLNPAQSYTAEPITFKVEGKYVTI	372
Db	291	QRPNESELEPTLPV-----MLDG--QEVPEVPSSELEPALPELMBELDGOEVEV	338
QY	373	-----IDGKQI--ENPNKE--IVEPYSYEAVYNDPEEFSVLITQNVAKRYYA	414
Db	339	PSELEPALPELMBELDGOEVEKPSVDLPEDR-----	373
QY	415	KNKNGSOVVYCFNADLKSPPDSEDGKTMTPDFTTGEVXY--THIAGRDLFKYTVAPRDT	473
Db	374	-----YEFNNKQDPSLAGE-----SETEYEITFEVYGNQ-----QNPVDI	407
QY	474	DPDTFLKHIKVYIEKGYREKQALIEYSGLTQYLRATQALAYITPDASFLDKLDYH	533
Db	408	DK-----KLPNRTGF--SGNMVETEDTKPEPGLMGQO-----SESVETKDTQTGMS	452
QY	534	GFGDMNDSTLVAKILVEYAADSNPP-----QLTDLDFEIPNNKKQOSLIGTQWMBEL	587
Db	453	G-----QTTPEVE-----TEDTKPEPGLMGQSESVETKDTQTGMSGQTASQVETED	501
QY	568	VDI--IRMEDKKEVLPVTHNLTKRTVYGLAGDRKDKHFELELKNNOELLSSQVTKDKT	646
Db	502	KEPGLVLMGQSESVET-----KDTQTGMSGQTPR-----QVETEDTKPEPGLMGQSE--	550
QY	647	NLEF--KDKAKATINKHGSESLTQGLPEGYSYLVEYETDSEGYKRVKVSQEVANATVSKTGI	705
Db	551	SVEFTKDTQTGMS--GFSEYTVTIVEDRPLVYHPHFNNEKPEVNEKREKPT-----	598
QY	706	TSDETLAEFNKKEPVPR--TGVDQKQINGVLLALIVYA	739
Db	599	-----KNITPLPATGDIENVLALITLILIS	624

RESULT 3  
T30856  
Protein F2 - Streptococcus pyogenes  
C.Species: Streptococcus pyogenes  
C.Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 11-May-2000  
C.Accession: T30856  
R.Jaffe, J.; Natanson-Yaron, S.; Caparon, M.G.; Hanski, E.  
Mol. Microbiol. 21, 373-384, 1996  
A.Title: Protein F2, a novel fibronectin-binding protein from Streptococcus pyogenes  
A.Reference number: Z20907; MUID:97011581  
A.Accession: T30856  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-1039 <JAF>  
A.Cross-references: EMBL:U031980; NID:g1654115; PID:g1654116; PIDN:AMC44522.1  
C.Genetics:  
A.Note: PRF2

```

Query Match          4.2%; Score 164; DB 2; Length 1039;
Best Local Similarity 19.7%; Pred. No. 0.087;
Matches 171; Conservative 124; Mismatches 279; Indels 296; Gaps 46;

QY 57 STPNAINP-DSSSESYRMWGYESYVGHPEYKQRYAHDLRNLGSSRSYQYVCNMLKRAF 115
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 110 NSPLAIGRYDAKTHQLUTFTFNLYING--LDRVQLSAETSLPLENKEYILE----- 156

QY 116 PLGSDSSVKKKKTKKHDGISTKFEDYAMSFRITGDDLNOK--LRAYMYGHONANGIMEG 173
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 157 -----NTNISDKRST--IGGGEITYKGTIVNLYVNGESTESNYTING 196

QY 174 L-----EPLNAIRVTOEAVYVYSDNAPISN-----PDESEKREESNVLVSTS 215
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 LSNVGGSTESINT--ETGEFWYVYVYVNNRRTIIPAVLVNLGCFARKTQGGENDSNLSVSA 254

```

[illegible]

A:Gene: clua  
 F:1-43/Domain: signal sequence #status predicted <SIG>  
 F:44-1243/Product: sex factor aggregation protein #status predicted <MAT>  
 Query Match 4.0%; Score 159; DB 2; Length 1243;  
 Best Local Similarity 19.6%; Pred. No. 0.23; Mismatches 243; Indels 314; Gaps 42;  
 Matches 160; Conservative 100;

QY	98	LEGSRSYQYCEFLKRAFLPGSDSSVKWKYKKRKHGISTKPFEDAMSFRITGDELINOKLRA	157
DB	335	MAGSLNYS-----KKTETBGGILPTISDGNQHSVS--FDN-----TL	370
QY	158	VMYNGHPQANGIMEGLEPLNLAIRVTOEALVWYISDNAPISNPDESEKRESESNLVSTSQL	217
DB	371	VVGSYEPSSGCVHVSQ--RPTSGANATGDS--WSSNPPTNETWASATAYIDYKAIGPS--	424
QY	218	SLMRQLKQLIDPNLNTK-----MPKQVPPDPOLSTFESDEKDKYKQNL--LSGL	270
DB	425	-----LDVTEWDVGTGKMSWGAMNLRPKDGGTSTISVTGTGD-----ANMALLINGQ	471
QY	271	VPTKPTPDPDPMPMPQOPQTTSVLRKKAIGDYSKLLEGATLDTGDNVNSFQARVSSN	330
DB	472	LENKTIPT--EPRIIPKPRATYYDDATPTDNTK-----AVTQTDGIDLN-----	516
QY	331	DIGERIELSD-GTYTLTEINSPAGYSIAE-----PITFKVEAGKVYTI-----ID	374
DB	517	--GALVNKQETERNWVLSNEVLPAGHEVIKSYVMTDPLPEGFKILDEQSKTILSPDYDLTFD	574
QY	375	GK-----QIENPKKEIYEPYSVEKYNDFEESVLTTONYAKFYAKKNGSSQV	423
DB	575	EKTNTVTLTAAYKATLEAMKDLQAOVPRK-----TLQGVY-----TKDSSS--	617
QY	424	YVCENADLKSPDSEDEGGTMTPEDT--TGEEVYTHIAGRDLEFKYVYKPRDTPDRTLKH	481
DB	618	---FKKDL-----ETLINDITYNSNEVE-----VHTDPRPREKSNEN	651
QY	482	IKKVIKGYRERKQALIE-----YSGLETFQ--LRATOLAIIYFTDSA-	522
DB	652	-----ASGTTINGQGIDVANATNTYKLLMDLSGKGAIASSKEDIVRG-----FYVDAAP	700
QY	523	-----ELDKDKLKDHYG-----FGDMNDSTLAVAKIYE-----	551
DB	701	DVAVDVLKNIYSKDSQGEKYGITAKVYSSVKAPAEVQVFLADAKIAPKGFQVYFVSVD	760
QY	552	-----YKODSNPQLT-----DLDFEIPNNKYYQ-----SLI-----	578
DB	761	PQTFYTNVYQGTGNNVETITQPMTFKEGASGAYQNTDIQIDEGNSIEGDTYKNNIVYPRKYK	820
QY	579	-----GTOMH--PEDLVDIIRMEDKKEVIFVTNLTLRKTVTG-----	614
DB	821	QYSVDSGKRWHSKDLRPTDSDSNYDYKLDNFETANGDYTKILLDGNFESSQWTLAKAKYT	880
QY	615	-----LAG-----DRTKDFHEIEELKNNKQELLISQYVTKDTN-----	647
DB	881	DKDGDNDIAQGFVLNASKGVDYTKDFENNHFQKGEKEVLIITFPKISDITSLASNDP	940
QY	648	-----LEFKGKATINLKHGESLFLQGL--PEGSYLVKLEPDSGEYKVKVNSQEVANA	698
DB	941	DLRLITITMSFKD-----VTLKGGTGALNAYL-----DKESKIYAPNIGOL--D	982
QY	699	IYVSKGITSDET--LAFENNKEPVNPFGVDQKINGLY	733
DB	983	TTSRT--VTGDNTKDKKITKSNTVTVKIPQRLPMLNKTV	1018

RESULT 5  
 F82884  
 hypothetical protein UU495 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: F82884  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: F82884

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-5005 <GIA>

A:Cross-references: GB:AE002148; GB:AF222894; NID:96899495; PID:MAF30907.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

A:Genetics:

A:Gene: UU495

A:Genetic code: SCC3

Query Match 4.0%; Score 157.5; DB 2; Length 5005;  
Best Local Similarity 19.8%; Pred. No. 2.5;  
Matches 158; Conservative 100; Mismatches 285; Indels 253; Gaps. 36;

```
QY 5 RPPNKLNTLNTQORVLSKNSKREF-----VTL-----VGVEFLMIFALVT 42
DB 1598 RIKGQNIISIPATISDDDKRYTKCMATNLGLDYEIVSVIKKQTKNVSPIIFELPS 1657
QY 43 SMVAKT--VGLVESSTPNAINDSSSEYRWYGESYVRGHPYTKQPRVAHDLRVNIE 99
DB 1658 GPICGYTLAPYFKITSLSK-----DPYHKKNTADUKINLR 1694
QY 100 GSRSGYVCFNKKAPPLGSDSSVKKWK-KHDGISTKPEDYAMSPRTGDELNOKLRAY 158
DB 1695 -IENIGASLFI-----KDIKIFRKNDGKOISIHKVT----- 1728
QY 159 MYNGHPONANGIMEGLEPLNAIRYQEAWYYSDNAIPISNPDESEFKRESESNLVTSQLS 218
DB 1729 -----ANANYPEMERKDLLRNEYTLERVY-----LKNKPFNGSVSES----- 1768
QY 219 LMRQALKOLIDNRLATKMPKQYPPDFOLSFESSEDKGKYNKGQYLLSGGLVPTPKPT 278
DB 1769 -----DYDLMIYENLNKTFK-----LLPTPKLGI 1793
QY 279 GPPMPNPOTTSVYLIRKVAIGDYSKL-EGATLQLTGDNVNSFOARVSSNDIGE--- 334
DB 1794 IGAPIKFISDNGAKVOL-KFAINDPDDVJAKENOTPKF---NIOPNNANLNEISEHGG 1849
QY 335 RIELSDG-TYTLTSLNSPAGYSIAEPIFEKVEAGKYTIIDGKQIENPKKEIVE-DYSVE 392
DB 1850 KYEIIDGKFEYANLN-----IKVKKYKY--KIYDEMQDVGNGYKIN 1894
QY 393 AYNDPEE-----FSVLTQNTA---KYYAKNKGSSQYVCFNADJLSPDSDEGK 442
DB 1895 FKNDYKEPNNVYDASVNTQTYVFNKFAIASFSNNLTDVANKONISINDS---R 1950
QY 443 TMRPPTGGEVYTHIAGDLPFKYTVKPRDTPDFTLKHKKVIEKGYREKGOAIEYSGI 502
DB 1951 VETIQIHFRAKATISDNKRYVNTNTI-PAPT-----IGNKNNIALNFE-L 1995
QY 503 TETQL---RAATQALAIYFETDSALDKLKDYGFGDMNDSTLAVAKILVEYAQDSNPP 559
DB 1996 NQNLISNLTYFAALY--SKDVNVDE-----NHANVYIKNVNPOITSTRK 2042
QY 560 QLTDLDFPIPNNNKYOSLGTOMHPEDLDIIMEDKKEIYI-----THNLTLR 610
DB 2043 SSYYVYDLKQOMDENKITISLLHSNDQI---FEDKNHMLKIAKISIDELAHDOITNS 2098
QY 611 YTWGLAGDRKIDHFETELKNNKQELLQTVKTDNTLFEKDGKATINLK----- 660
DB 2099 T-----THDIDLLEKENNEMMLKTQILN-KERTKYRVKVVFTSKPSDTIYNGVN 2149
QY 661 -----HGESLTLOGLEPGYSYLVKETDSEGYKVKVNSQEVANAV--SKTGITSD 708
DB 2150 KQIIVYPINNHNTNIDLKTEKSTLNSVKPT-----KTNKKNESIKLMGFWKGTSGSLE 2204
QY 709 ETILA---FENKPEV 720
DB 2205 NKYAKLVYKDNNOOI 2220
```

## RESULT 6

701172

toxin B - Clostridium difficile

C:Species: Clostridium difficile

C:Date: 12-Feb-1998 #sequence\_revision 20-Feb-1998 #text\_change 15-Oct-1999

C:Accession: S701172; S44271

R:yon Eichel-Streiber, C.; Meyer zu Heringdorf, D.; Habermann, E.; Sartingen, S.

Mol. Microbiol. 17, 313-321, 1995

A:Title: Cloning in on the toxic domain through analysis of a variant Clostridium dif

A:Reference number: S701172; MUID:96079281

A:Accession: S701172

A:Status: Preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-2367 <ON>

A:Cross-references: EMBL:Z23277; NID:9761713; PID:CAAB0815.1; PID:9761714

A:Experimental source: isolate 1470

R:Sartingen, S.; von Eichel-Streiber, C.

submitted to the EMBL Data Library, July 1993

A:Description: Cloning and sequencing of an Clostridium difficile toxin B mutant.

A:Reference number: S44271

A:Accession: S44271

A:Molecule type: DNA

A:Residues: 1-1323, 'N', 1325-2367 <SAR>

A:Cross-references: EMBL:Z23277

C:Superfamily: cpl repeat homology

C:Keywords: cytotoxin

Query Match 3.8%; Score 151; DB 2; Length 2367;  
Best Local Similarity 19.9%; Pred. No. 1.9;  
Matches 171; Conservative 94; Mismatches 292; Indels 304; Gaps 42;

```
QY 5 RPPNKLNTLNTQORVLSKNSKREFYTLVGVFLMIFALVTSNMGAKYVGLVE--SS 57
DB 1568 KPMNRKGSNTSDSLMSFLSESNIKSIFVNFLOSNIKFILDANFLISGTTSGGEFICD 1627
QY 58 TPNAINP-----DSSSEYRWY-GYESYVRGHPYKQPRVAHDLAVNLEGSSTOYCFN 110
DB 1628 ENNNIQPYFIKFNTLETNTLYVGNRQNMIVEPNY-----DIDDSGDISSTVIN 1676
QY 111 LKAPPLGSDSSVKKWKYKHHDSISTKFEYAMSPRTGDELNOKLRAYVNGHPQ-----N 166
DB 1677 FSQKLYLYGIDSVNK-----VVISPNLYDEIINTYVEENNTPEYIYLD 1722
QY 167 ANGIMEGLEPLNAIRYQEAWYYSDNAIPISNPDESEFKRESESNLVTSQLSLMRQALKQ 226
DB 1723 ANIYNEKIN-VAINDLISIRYV-----SNDGNF-----ILMSTSEKKSQVKIR 1767
QY 227 LI---DPNLAITKMPKQYPPDFOLSFESSEDKG-----KYNKYQNLISG-- 268
DB 1768 FVNVFVKDKTLANKLS-----FNFSKQDVPVSEIILSFTPSY---YEDGLIGD 1813
QY 269 -GLVPTPKPTPGDPRMPNPOTTSVYLIRKVAIGDYSKLLEGATLQLTGDNVNSFOARV 327
DB 1814 LGLV-----SLYNEKFTYNNGMWVSG--LIYINDSLYKRPV- 1850
QY 328 SSNDIGERIELSDGYTLTSLNSPAGYSIAEPI-----TKVEAGKY 370
DB 1851 -NNLITGFYVGDKKYFNPINGAA-SIGETIIDDKNYFNQSGVLQGVSTEDGFKY 1908
QY 371 ---TI---IDGKQIENPKKEIVEP---YSYEAIVDFEFSVLTQNTQ-----AKFY 413
DB 1909 FAPANTIDENLEGEALIDFGKLIIDENIYFEDNYRGAAVEKELDEMYFSPETGAKR 1968
QY 414 AKKNGSSQVYVCFNAD--LKSPDSEDDGKTMPTFTGGEVYTHIAG----- 461
DB 1969 GLNQIGDCK--YFNSDGVQKGFVSINDKHYFEDSGYWKAGYTLIDGKHFFAENGEM 2026
QY 462 -----DLFKYTVKPRDTPDFTLKHKKVIEKGYREKGOAIEYSGLTQLRAATQL 513
DB 2027 QIGVFNTEDGFKI-----FAHNHEDL-----GNEGEISYSGILNFNKK----- 2066
```

```

QY 514 AIYFTDSAE-----LDKDKLKHGFGDMN-----DSTLAVAKILV 550
    ||||| |||
Db 2067 -IYFDSFTAVVGMKDLDEGSKYFDEDAEYIGLSLINDGOYFENDGIMQVGFVTI 2125
QY 551 E-----YAODSNPQLTDLID-----FFIPNNN-----572
    ||| |||
Db 2126 NDKVFFYSDGIIIESGQNIIDNMFYIDDNGIYQIGVFEDISDGKYYAFAPANTYNDNIYQ 2185
QY 573 --KYOSL-----IGTQHPEDLVDIIRMEDKKKEVIVPHTNLTKRTVTGL 615
    ||| |||
Db 2186 AVEYSGIYFVGEDEVYFGETYTIETGW-----IYDMENESDKYFVEPET-----KKACKGI 2236
QY 616 AGRTKDFHE-----IELKNNKQELLSQTVKTDKNLEFKOKATINLK---HGE 663
    ||| |||
Db 2237 NLIDIKYFDEKIMGTGLISENNNYF-----NENGELQFGYINIEDKMEYFGE 2288
QY 664 SLTLQ---GLPEGYSLYKE 680
    ||| |||
Db 2289 DGYMGIQVFTPDGFKYFAHQ 2309

```

```

RESULT 7
T28679
fibrinogen-binding protein homolog - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-May-2000
C:Accession: T28679
R:Jorgensen, E.; McCrea, K.; N. Eidin, D.; O'Connell, D.; Cox, J.; Hook, M.; Foster, T.
Microbiology 144, 3387-3395, 1998
A:Title: Three new members of the serine-aspartate repeat protein multigene family of St
A:Reference number: Z20510; MID:9908700
A:Accession: T28679
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1315 <JOS>
A:Cross-references: EMBL:AJ005646; NID:el318791; PID:el318792; PIDN:CAA06651.1
C:Genetics:
A:gene: sdfr

```

```

Query Match 3.8%; Score 149; DB 2; Length 1315;
Best Local Similarity 20.2%; Pred. No. 0.98;
Matches 203; Conservative 93; Mismatches 336; Indels 372; Gaps 45;
QY 8 NKLNTINTORVLSKNS--KRFVTLVGVFLMIFALVTSWGAKTGVLVESSTP-----59
    ||| |||
Db 208 NSNNENNADILIKRSTAPKRLNR-----MRIAQVPSSTREAKNVNDLITSNTLTIVVA 262
QY 60 ---NAINP--DSSEYRWYGVESYVGHPIYKQFRAHDLRVNLEGSRSYQYCFNLKKA 114
    ||| |||
Db 263 DKNKKIVPADDYLSLKSQITVDKVKSGDYF-----TIKYSQVQVYGLNPEDI 311
QY 115 FPLGS---DSYKMKWKYKHDG---ISTKFEDYAMSPRTIGDELNOKLAAYVNGH---163
    ||| |||
Db 312 KNIGDIKDPNNGETIATAKHDITANNLITTYFTYV-----DRENSVQKINTYITMDA 364
QY 164 ---PONANGIMEGLEPNAIRYQEAVMYSDNAPISNPDESFEKRESSENL-----211
    ||| |||
Db 365 DTIPVSKNDVEFNVTIGNTTNTTANIQDPDYVANEKNSIGSAFTETVSHVAKENPGY 424
QY 212 -----VSTQSLSLMRQALK-----QLIDPNLATMPKQVPPDFQSLSTFESDGDKDYKGY 262
    ||| |||
Db 425 KQITVYVPSNSLTNAKLQAYHSSYPNNIGQINKDVT---IKIYQ--VPKGYTLNKG 480
QY 263 ONLLSGLVTKPPTGDPMPRPNQPTSVLIRKVIAGYSK-----305
    ||| |||
Db 481 D-----VATKELT-----DYTNQYLOKITIYGGNNNAVIDEGNADSYVYVAVN 522
    ||| |||
QY 306 -----LLEGATLQLTGD-----NV 319
    ||| |||
Db 523 TKFOYTNSEPTLVQMATLSTGSKSVSTGNALGFTTNQSGAGQOEYKIKGNVYVWEDFNK 582

```

```

QY 320 NSFQ-----ARFSSN---DIGERIEISDGYVTLTEL-----NSPAGYS 355
    ||| |||
Db 583 NGVQELGEGKGVNVTYVFNNTNTKAYGEAVTKEDGSLYLPNLNPGDYRVFESNLRPGYE 642
QY 356 IAE-----358
    |||
Db 643 VTPSKQGNNEELDSNGLSVITVNGKDNLSADLGIYKPKYVNLGDYVWEDTINKNGIDQDE 702
QY 359 -----PIYFKVAGKY-----IIDGK---QIENPKEL-----VPPYVEAVANDF 397
    ||| |||
Db 703 KGISGVVTLTKDERNGVNLKVTITDADGKYKFTDLDNNGNRYVEFTTPEGYPTTVTSGSDI 762
QY 398 EEPF-VLTQNY-----AKFYAKKNGSSQVYVCFNLDLSPSPSED---GGKT 443
    ||| |||
Db 763 EKDSNGLTITGVINGADNMFLDSGFYKTPRYNLGNVYWEPTNNDGKO--DSTKKGISGV 820
QY 444 MTPDFTTGEVYKTYHAGRDLFYKYVYKPRDTPDTFLKIKVIEKGYRKGQAIEY---499
    ||| |||
Db 821 VTLKNGEVLQTTKTDKDKGYFT-----GLENGYKVEEETP 859
QY 500 SGLTETQLRAATOLAI-YITDSAEIDDK-----LKDY-----HGFG 536
    ||| |||
Db 860 SGYTPYQVSGTDEGIDSNSTGTGVIKDKNDPTIDSGFYKPYNLGDYVWEDTNKNGVQ 919
QY 537 DMNDSTLAVA-----KIIVEYAQDSNP-QLPDLDFEIPNNKYOSSLIGTQWHPED 586
    ||| |||
Db 920 DKDEKGISGVVTLTKDENNDVLTQVTTDENGKYQFTDL-----NNGTYK-----963
QY 587 LVDIIRMEDKREYIP--VTHNLTLRK-----IYVGL---AGDRTKFHFIEELKNNQEL 636
    ||| |||
Db 964 -----VEEFTSGTTPSVTSGNTEKDSNGLTITGVYKADNMNTLDSGF---YKTPRYSL 1016
QY 637 -----LSQTVKTDKTNLEFKDGKATINLKHESLTLQGLPBGYSYLYKETDSEGYKVK 689
    ||| |||
Db 1017 GDYVWYDSNKGDKDSTPEKIKDYKVTLLNKEKEVIGTGTITDENGKCYCFNDLSGKYKVI 1076
QY 690 VNSQEVANATVSKTIGTISDETLAFENKKEPVYVGVQDKI-NGY 732
    ||| |||
Db 1077 F--EKPAGLTGTGTNTTEDDKDADGGEVDVITIDHDFTLIDNGY 1118

```

```

RESULT 8
G82875
hypothetical protein U0558 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82875
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: G82875
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1883 <GLA>
A:Cross-references: GB:AE002154; GB:AF222894; NID:96899557; PIDN:AAF30971.1; GSPDB:GN
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0558
A:Genetic code: SGC3
Query Match 3.8%; Score 148.5; DB 2; Length 1883;
Best Local Similarity 19.1%; Pred. No. 1.9;
Matches 155; Conservative 102; Mismatches 276; Indels 277; Gaps 34;
QY 103 SYQVYCNLKKAP--PLGSSSVKWKYKHKHDGISTKREDYAMS-----143
    ||| |||
Db 951 NFDTSMFNLKTYFDHPMTSNEKRKAERITDAMFKYDPDNISKSLKLDKANAASVPY 1010
QY 144 ----PRITGDELNOK-----LRAVYNGHPONANGIMEGLEPNAIRVTOEAVVYS 191
    ||| |||
Db 1011 KYVLSQITIGEDITNKTFDQKRYALTLIQYNAASKIMDKSKEEVSXLDPELI--EIKIYY-1067

```



A:Title: A Fibrinogen-binding protein of Staphylococcus epidermidis.

A:Reference number: Z20781; MUID:98261511

A:Accession: T30214

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1092 <NII>

A:Cross-references: EMBL:Y17116; NID:e1296734; PID:e1296735; PIDN:CAA76638.1

Query Match 3.7%; Score 144.5; DB 2; Length 1092;  
Best Local Similarity 18.0%; Pred. No. 1.4;  
Matches 159; Conservative 145; Mismatches 269; Indels 309; Gaps 44;

```
QY 8 NKTINTLTQR-VLSKNSK-----RTVTLVGVLMIFALVTSVAKTYVGL-----V 54
DB 3 NKNKNLTTRKKPLANKSMYAKRFTVGTASI-----VIGATLLEGLGHNEAKAE 52
QY 55 ESSTPAINDSSSEXYRWYGESEYVGHHPYKQFVAHDLRVNLGSRXYQVCFNLKA 114
DB 53 ENSVQVKKDSNTDLESD---DSNDQSDDEKNDVYINNQSINTDNNQI-----IK-- 100
QY 115 FPLGSDSVKKWKHKHDGISTKFEDYAMS-----PRITGDELNQKLRAVWYNG 162
DB 101 -----KEETNNYDGLIEKRSEDETESTYNDENEATFLAQTPQDNTHLTREEVKESS 151
QY 163 HPQNGIMEGLEPLAINTQEAIVYVYGDMAISN-PDESEKRESESNLYSTQSLMR 221
DB 152 SVSSNSSTIDTAQOPSHTTTNRRESVQTSNDVEDSHVSDPANSKIKESNTSEKENTTE 211
QY 222 QALKQIDPLNATKMKQVDDPQLSIFESDQKDKRYKNGYQNLSSGLVPTKPTPGDP 281
DB 212 Q-----PNKV-----KEDSTTSQPSGYTNI-----DE 233
QY 282 PMPNPQPTTSLIRKYAIGDYSKLEGATLQLTGDNVNSFQARVSSNDIGERIELSDG 341
DB 234 KI-SNDELINLPINEYE--NKARPLSTSAOPSIKRYVYVQLAQSGSNVNLIRKVTQD 290
QY 342 TYVLTFLNSPAGSIAB-----PIFE-----VEAGKYVTI-IDGKQI----- 378
DB 291 SIT-----EGTDSGEYKAKHDAENLIYDVEYDVKSGDTMTVDKKTVPISDLT 343
QY 379 -----ENPKIEVEPYS--VEAYNDEFEFSVLQNTAKKYYAK 415
DB 344 DSFTPIKIDNSELITATGIDYDKNNKQIYTTFDYDKYENIKAHKLKTS-----YIDK 397
QY 416 NK--NGSSQVYVCENADLSPDSEDGKTMPPDF-----TTGEVK--YTHIAGDLEK 465
DB 398 SKVPNNNTKLDVETKALSSV-----NKTIVEYQRPNENRTANLQSMPTNI--DTKN 448
QY 466 YVKKPDDTDPDTFLKHKIKYIEKGYREKQALIEYSGLTE---TQLRAAQOLAIYFTTDSA 522
DB 449 HNY-----EQTIIYINPLR-----YSAKETNWNISGNGEGSTIIIDSTLIKRYKVDNQ 497
QY 523 EL-DKDKLDYHFGDMNSTLAVAKILVEYVADSNPQLTDLDFIPNNKKYQSLIGTQ 581
DB 498 NLPDSKRITDISEYEYVND-----DYAO-----LGNNDVYNINFGNI 535
QY 582 WHEDELVDIIRMED--KKEVIPTVNLTLKRYVTGLAGD----- 618
DB 536 DSBPY-IIVKYSKDPKADYDTTIOQTVMQTTINEYTGERTASYONTIAFTSSGGGQG 594
QY 619 -----RTKDPHEF-----TELKNNKQELL-----SQYKTKTKLNEK 651
DB 595 DLPEKTYTKIGDVVEDVDKDGIONTNDNEKPLSNVLVTLVYPDGTSKSYRTD-----E 648
QY 652 DGKATIN-LKHG--ESLFTLOGLEPGSYLVKET-----DSEGYKQVNSQEVANATV-- 700
DB 649 DGKTYQDGLAKNGLTLYKITEP-TPEGITPLKHSGTNPALDSBGSNWSVYTIINGODDMTIDS 707
QY 701 -----SKTGITTSDE-----TLAFEN 715
DB 708 GFYQTPKYSLSGNVWYDTNKKDGIQGDDEKGISGVKVTLDEN 749
```

RESULT 11

A42404 collagen adhesin - Staphylococcus aureus

C:Species: Staphylococcus aureus

C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Nov-1995

C:Accession: A42404; 527665

R:Patil, J.M.; Jonsson, H.; Guss, B.; Switalski, L.M.; Wlberg, K.; Lindberg, M.; Hook

J. Biol. Chem. 267, 4766-4772, 1992

A:Title: Molecular characterization and expression of a gene encoding a Staphylococcus

A:Reference number: A42404; MUID:92165839

A:Contents: FDI 574

A:Accession: A42404

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1185 <PAT>

A:Cross-references: EMBL:M81736

A:Note: sequence extracted from NCBI backbone (NCBIP:83982)

Query Match 3.6%; Score 143; DB 2; Length 1185;  
Best Local Similarity 19.7%; Pred. No. 1.9;  
Matches 151; Conservative 108; Mismatches 279; Indels 230; Gaps 40;

```
QY 99 EGSRSYQVCFNLK-----KAPPLGSDSVKKWKHKHDGISTKFEDYAMSPLITGDE 150
DB 269 QGYGTSNYSINRYKTKITNEOKEFVNNSOA---WYQEHG-----KEEYNGAS 313
QY 151 LNOKLRAVWYNGHPQNGANGIMEG-----LEPLNAIRVTOEAVYVYSDNAPISNPDESFK- 204
DB 314 FNHTV-----HNINANGIEGTVAKELVKQKDDTK-----APIAN--VKFKL 355
QY 205 RESESNLYSTQSLMRQALKQIDPLNATKMKQVDDPQLSIFSE-----DKGDKY- 258
DB 356 SKKDGSVYVDMQ-----KEIEITTDANGIANIKALPSGDYTLKEIAPRPTEDKDKREY 410
QY 259 -----NKGYNLSSGL-----VPTKPTPGDPMPNPQPTTSLIRKYAIGDS 304
DB 411 FYMKDTDNQGYFTTIEENAKAIEKTKDVSQAKVWEGTQKAP-----TIYFKLYKDDNDQ 464
QY 305 -----KLEGATLQLTGDNVNSFQARVSSND-IGERIELSDGYTLTELN----- 349
DB 465 NTTPVDKAEIKLEDDGTTKYTWMSNL-----PENDKNGKALK-----YLVEVANAQGE 511
QY 350 --SPAGSIAB-----PI-TRKVEAGKYTTIIDGQIENPKYIEVYSVAIYND 396
DB 512 DTPPEGYTKKENGLVYVNTKEKPIETTSISEGKYWDKDKNDODGKRPER-----VSANLILAN 566
QY 397 FEEFSVLTTQNTAKFY-----AKNNGSSQVYVCENADLSPDSEDGKTMPTPTTG 451
DB 567 GERVKTLDVISTENMKYERKDLPRYDEG--KIEEYVTEHVKDYTTDINGTTITNKYTPG 625
QY 452 EVKYTHIAGHDLFEKYVYKPPDDTDPDTFLKHKIKYIEK-----GYREKG 494
DB 626 ETSATYTKMMDNNDGKRPTIEIKVELYODGKATGKTALINSNMWTHWTGIDLEKAG 685
QY 495 QALEYSGLTETQIARAQO-----LAIYFTDSAELODKLKYDHGFGDMNDSTL 543
DB 686 QQVKYVVEELTKVKGTYTHVNDNDMGNLIVTNKYTPETTSISEGKYWD-----DKDN--- 737
QY 544 AVAKILVEYVADSNPQLTDLDFEIPNNKKYQSL-----IGQWHPDELVDIIRMEDKKEVI 600
DB 738 -----QDGKREKVSVN--LADGEKVTLDVTSFTNMKYE--FKDLPKYDECKKI- 784
QY 601 PYTHNLTKRYTVGLAGDRTKDPHEFE-----LKNKKQELLSQLYKT--DKTNLE----- 649
DB 785 -----EYTVY--EDHVKQDVTIDINGTTITNKKTTPGTSATYVAKNMDNNDGKRPR 833
QY 650 -----FRDGKAT-----INLKHGSLTLQGLPE-----GYSYLKVE--TDSEGYKQVKN 691
DB 834 TEIKVELYODGKATGKTALINSNMWTHWTGIDLEKAGQOQVYVVEELTKVKGTYTHVD 893
QY 692 SQEVANATYVSK-----TGITSDETLAFENKKEPVYPTGVQDKI--NG 731
```



Db 894 NNDMGNLIVTNKYPETTSISGEKVDKDNQDGKRPEKVSYNLLANG 941

RESULT 12  
T03821

P:olipoprotein p65 - Mycoplasma hyopneumoniae  
 N:Alternate names: major immunogenic surface lipoprotein  
 C:Species: Mycoplasma hyopneumoniae  
 C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 07-Dec-1999  
 C:Accession: T03821  
 R:Heidari, M.B.; Kim, M.F.; McIntosh, M.A.; Wise, K.W.  
 submitted to the EMBL Data Library, July 1997  
 A:Reference number: Z15102  
 A:Accession: T03821  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-627 <HE1>  
 A:Cross-references: EMBL:AF013714; NID:g2338561; PIDN:AAB6173.1; PID:g23385656  
 C:Genetics:  
 A:Genetic code: SGC3

Query Match	3.6%	Score 142.5;	DB 2,	Length 627;
Best Local Similarity	20.5%	Pred. No. 0.74;		
Matches 163; Conservative	116;	Mismatches	265;	Indels 251;
				Gaps 41.

QY	19	LSLSKRETVYLVGVEFLMFLMFLVYNSMGATKVTFCVLVSESSPTPMANPDOSSEKRYNYGESY	78
Db	1	MKKARKK-LWLSLTLTLPAPSVFTTLLISAC--LONNSLSTVEYNALDSDSTAGNEE	56
QY	79	VRGHPYKQFRAVDLRVNLNIGSRKSOYV-CFNLKAAPLGSOSVKKWKKHGIS--	134
Db	57	T-----YRDPGTLDKOGNLSG--QSYPAFYAYLQK--LNKNSLV-----SYDNLAISG	102
QY	135	-----TFEDYAMS--PRIT-----GDELNOQLRAVMINGHPQANMIGETLEPL	177
Db	103	TTTENMUYLNPRTYPPGKMSDNPLVYNSGNEKEYNIGSVFGDFNKDSYGLVEKKKA	162
QY	178	NAIRVTOEAVMYSDNMPISNPDSFKRESSESNVSTISOLIMQALKOLIDPLATKMP	237
Db	163	NLMMSGA-----NDPFLAIFNEFKMA-----SITPKSSEAKKLIDP--ERA	206
QY	238	KQVDDPOLSLFSESDGDKYKNKYONL-----LSGGLVPTKPTPGDPMPPNQ	287
Db	207	NFLAEKGMLTKAEVKNKKEETINTMDNLIKELKALNPKLSINLYGKTLPMNSG-----	258
QY	288	PQTSVVLIRKAIIDYSKLEGGALLOLTGDNVNSFQARVSSN-----DIGERELSDGT	342
Db	259	-----FIFILKYLTLTYVK-IETDIFINIEPEKINKIIRETAKKNVNYIDVYDSIMWSD	313
QY	343	YTLTELN-----SPAGY-SIAEPTTFVEAGKYTIIDGKOIENPKIEVPEASVEAND	366
Db	314	KNLMANKNDFHPHSIGQIKKIAHOLLKL-----TLDQEKSDSNAE--ELKNTTFED	364
QY	397	FEFSPVLTQNYAKFYAAKKNKNGSOVAYCFNMADLKSPPSEDDGKTMPTDFTTGEVYKT	456
Db	365	FDE-----NKPTYSKYV-----	376
QY	457	HIAORDLEKYVRKRDNDPDLFLKHKKIYVEKGRREGQALIEYSGLTEQTQRAATOLAIY	516
Db	377	-----DLSVFAKSNKEPELEK-LNEKKQTSSEFLAQKST-----	407
QY	517	YETTSAELEDKDLKDYHGFGDMNSTLAVAKILVEYADSNSPOLTDLP--FIPNNKY	574
Db	408	FDTDOEPAIKD--DKRTFGUI-----VREIV-----SLPFDNPFDRRELIPYKNPF	451
QY	575	OSLIGTQWHPED--LVDIIRMEDK-----EVIPTVHNLTURKRYTVGLAGDRTKD	622
Db	452	VKAIINSYLGKPAAGSLIKDIEOLENKKVADYARPNIKIFDTLISFIRKIMVAFPA-----E	506
QY	623	FHPFIELEKNNKQ--ELLSQYKT--DKTNLEFNDG--KATILMKGESITLQGLP	671

```

Db 507 INTDDEIEKEFFMSQOIIIFILIRLNAILSPFDITKL--KDSATFKILMLKPEOIIITLLGIG 564a
Qy 612 EGVSY-----LVKTESEGYKRVKVNSEEVANATVSTGTTGTTSETLAFENKPEPV 721a
Db 565 KTPSPVPRPEKPKDDGSMNQDTTSSQK-----QE-----SGTGT-DSTATTENQKPAE 612a
Qy 722 PT-----GVQKIN 730
Db 613 QTNSEOSTSDSKSN 627

```

RESULT 13

E82885  
hypothetical protein UU481 [imported] - Ureaplasma urealyticum  
C:Species: Ureaplasma urealyticum  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: E82885  
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Helner, C.R.; Chen, E.Y.; Cassell, G.H.  
submitted to GenBank, February 2000  
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a  
A:Reference number: A82870  
A:Accession: E82885  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-888 <GLA>  
A:Cross-references: GB:AF002145; GB:AF22894; NID:56899476; PIDN:AAF30893.1; GSPDB:GNP  
A:Experimental source: serovar 3; biovar 1  
C:Genetics:  
A:Gene: UU481  
A:Genetic code: SGC3

Query Match	3.5%	Score 139.5	DB 2	Length 888
Best Local Similarity	18.4%	Pred. No. 1.9		
Matches 150	Conservative 125	Mismatches 256	Indels 283	Gaps 39

```

QY 48 KTVLWVSSSPNMAINPDSSSEYRW-----YGESVYRGHPYKQORVAHDL-RV 96
Db 72 KRIDVLEVSASN-----KNSAKIKWNLTDLPETERYRTS-----AFKKNKPDIAKW 119
QY 97 NLEGRSYOVYCFNLKKAAPLGSDSVYKWKYKHDGISTKEEDYAMSP-----144
Db 120 MLIITKNNAVITGNNLDTFKTLVLDKILKLKLESNNDDIYNNQENAVNNRVIQITKAFFS 179
QY 145 RITGDELNOKLRAVWY-NGHPQ-----NAN 168
Db 180 RINSVYIKKKIKILVYEDNHNHOMETNELLSNNKTSYMFNDISKILTPNNKYSFKKILISNN 239
QY 169 GIMELEELNMAIRYQEAWVWYSDNAPISNDESPKRESEINLVSTQSLMRQALQOLI 228
Db 240 NIEE-ELLNSNNINSDSTYHHAIPA-----VFTEKGDEN-----OLW 276
QY 229 DPNLATKMPKPOVDFOFISIESEDK---GDKY-----NKGYNLISGLVLPKRPPTPG 279
Db 277 KSNIGITIAK-----FRGL-ESLDKILENDKILVYFANKKYPN-----315
QY 280 DPMPPNPOPTTSVILIRKAYAGDYSKLELGATLQLTGDNVNSFOARVPSSNDIGERIELS 339
Db 316 -----NODMKMYTMSITDSR--N 332
QY 340 DGYT-----LTELNSPAGYSIAEPITFKVEAGKYTTIIDOKI-----ENPN 382
Db 333 DGLVANIKNHDIWNLDEGESEYIKK-IRMIKKPNNAVEEDVNGNKIISPINDAORRTLENTN 391
QY 383 -KEIPEYSVEAYNDFEE-----FSVLTTPON-----YAKFYAAKNKNGSS-----421
Db 392 FKTNAITAIYNAKKYGMWDPTINSAKYAFETITTDGTLNLTNSKFOLVFKVYSOSTKTFI 451
QY 422 QVYCYCFNADLKSPPDSEDEGKTMTPDFTTGEKYYKTHLAGRDLKFYVYKPRDIDPDFELKH 481
Db 452 QFINLVKRDKNMILLEGESISNLEPNTEYILSTIKLSRPGNDLF-----DTELNTNINI 504
QY 482 IK---KXIEGKRGKGOAIEYSGITETQLRAAQOLAIIYTTSDSAELDKDKLAKYHGFCD 537

```



```

Db      505 I EONOKIILTKGIDPIESTETTKTNDLRNR-----IIRKID---IKD 546
Oy      538 MNDSTLAVAKIIVEAADSNPQ-----LQ-----DLDFPIPN--NNKY--OSLIGTOW 582
Db      547 YNDG-----RLVLEYRNNNSTSKREVSLLTNKQEDYSVNIENULTYNRREYTLNIRKIN 602
Oy      583 HPBDL--VDIIRNEDKKEYI-PTHNLTLRKTYTGLAGRTKPFHEIELKNNKQELLSQ 639
Db      603 NTNNISIDISKLNDKTFVVEPSVTKLTNFKNTNF-----DEKKQIWSR 647
Oy      640 TVKTDNTNLEFKDKGKATINLKHGESLTLQGLPGYSYLVEKTESEGKYKVNQGEVANAT 699
Db      648 IMFNTNTNDEMFSSQLN-----STKIRAIYVNLN--TKIDNPITIKI--VNDNGKPKLO 700
Oy      700 VSKTGITSETLAFENNKKEPVVPTGVOKINGYL 733
Db      701 IHINININD-----QNNYEKRLRVKFNKPNKSYL 730

```

## RESULT 14

F71909  
hypothetical protein jhp0613 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 08-Oct-1999  
C:Accession: F71909  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A:Reference number: A71800; M0ID:99120557  
A:Accession: F71909  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1167 <ARN>  
A:Cross-references: GB:AE001493; GB:AE001439; NID:g4155161; PIDN:AAD06194.1; PID:g415516  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: jhp0613

```

Query Match      3.5%; Score 139.5; DB 2; Length 1167;
Best Local Similarity 19.1%; Pred. No. 3;
Matches 136; Conservative 110; Mismatches 282; Indels 185; Gaps 32;

Oy      103 SYGVYCFNLKAPPLGSDSSVKKM-----YKKHD-----GISYKFDYAMSPR 145
Db      504 SYTIALINITQA-AQNRDSSLKFKNIALTDSLDYLEKKDKGVIFGFEYLFEDLKENKE 562
Oy      146 ITGDELINQKRAVYNGHPQANGIM--EGLEPLNAIRVTOEAVWYYSNADPISNDES 202
Db      563 IKRTMEQONIRVILI--GNPYSSGAKSENNDNONLSHPKLEKRYEYTGKNSIAQN---- 616
Oy      203 FKREESNLVSTQSLSIMRAKL-----QLIDPLMALKMPQVYDDQSLFEES 251
Db      617 -KSTRTDILLHS--IRNASDLAKDKGLGVNGVSFIDSADGFRRCVADPS-ELIYAL 672
Oy      252 EDGDGVYKNGYQMLSGGLVPTKPTPGDPPMPNQPOTSVL-----IKKYA 299
Db      673 NLKGNMARTSEER-----KKOGGCIIPSGSRVYAIILFFPKDKAPNHTIYYE 721
Oy      300 IGDYKLLLEGATLQLTGDNVNSFOARVSSNDIGERI-ELSDGTYYLTTELNSPAGYSIAE 358
Db      722 VEDYLKREAKLNTLANFENLDSVPKEITFPNDGDMINQNDQEKILPLKRDKKSIEN 781
Oy      359 PIPEKVEAGKVYITLIDGKQLENPKVEPYSEAYNDEEFPVLTQNTAKKYAKNKN 418
Db      782 AI-FDLNSNGVKTSRD-----PW---VYNFSQKTLMSQVN----- 813
Oy      419 GSSQVAVCFNADLK-----SPDSEEDGKGTMPDFTTGVEVKTTHAGRDLEKYYT 468
Db      814 ----CIDTYNADLKRENERREAFKORTAKDKIKKSAD-----RKHLNDEIITDTKT 863

```

```

Oy      469 KPRDTPDTFLKHKIKVTEGKREKGAIEYSGITETQL--RAATQALYY----FTDS 521
Db      864 KIAMTDG--LKN-KLI-----KNENLPESGMERVRLALYRFNKQWMLYKMLNINQ 912
Oy      522 AELDK--DK-----LKQYHGFQDMNDSTLAVAKIIVEAADSNPQDLDLDFIPNNKY 574
Db      913 SOLPKIIPDKSARNVYINTGVNGKDFSAVSPISDFSLIS-PNQAYPLYYDDDLGNRY 971
Oy      575 QSLIGTQWHEBEDVDIIRMEDKKEVIP-----VTHNLTLRKTYTGLAGRTKDF 623
Db      972 NALSGY-----ALNLFERRHYKDNALTEEEIIFYIAYAFHKGVLEKXKNSLAKENPR-- 1023
Oy      624 HPEIELKNNKQELLSQTVKTDKTNLEFKDKGKATINLKHG--ESLTLQGLEGSYLYKET 661
Db      1024 ---IALSEDFKELSMUKELAEHLNANESGEMHTSVKHNLESAEMEGY---YDVIOMKR 1077
Oy      682 DSEGYKKV-----VNSGEVANATYSKTGITSEDTLAFENN 716
Db      1078 DKRGDRIKYHHHTITTOIPKPAFDYVYNGKSAIDWYIERIQITKDKDSLIENN 1130

```

## RESULT 15

B64546  
chaperone and heat shock protein C62.5 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 20-Aug-1999  
C:Accession: B64546  
R:Tom, J.F.; White, O.; Keriavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKe  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; M0ID:97394467  
A:Accession: B64546  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-621 <TOM>  
A:Cross-references: GB:AE000541; GB:AE000511; NID:g2313299; PIDN:AAD07278.1; PID:g231  
C:Superfamily: heat shock protein 90

```

Query Match      3.5%; Score 139; DB 2; Length 621;
Best Local Similarity 21.9%; Pred. No. 1.2;
Matches 130; Conservative 85; Mismatches 227; Indels 152; Gaps 30;

Oy      201 ESKFRESEN-----LVSTQSLMRQALKQIDLPNATKMKQVDPDFOLSFESERKG- 255
Db      28 EILFRELIVSNASDALDKLNTLMLTDEKLK-----GLTFTPSIHLSFDSQKTYLTIKDNGI 82
Oy      256 --DKYN-----KGYONLLSGGLVPTKPTPGDPPMPNQPOTSVLIRKRYAIGDY 303
Db      83 GMDKNDLIEHLGLTIAKSGTINPLSA-----LSGD-----KKDSALIDGFGVGFY 127
Oy      304 SKLEGATLQLTGDNVNSFOARVPSSN-----DIGERIEISDGTYYLTTELNSPAGY---- 354
Db      128 SAEVAVASKIYVQTKVNSDAVAVMSDGKGFSEISCEVKDEQGETELTLFLKDEDSHPASR 187
Oy      355 -----STAEPTTFVEAGKYITIIDGK-QIENPKVEIPEYSEAYNDFEPEF----- 400
Db      188 WEIDSVYKATSEHIPPFI-----FLTYTDTYHNEGQDQKKEIKBEKQCIQASALMKMK 243
Oy      401 SVLTQNVAKFYAKKNGSSQVAVCFNADLKSPDSEEDGKGTMPDFTTGVEKYYTHI-- 458
Db      244 SELKDKRYKFEYQSFANDNEPLSYINK-----VEGSLLEYTLEFY 284
Oy      459 ----AGRDLEK-----YTVKPRDTPD-----TFLKHKKVIEKG-----YREKG 494
Db      285 IPSTAPDMERVYKSGGVKLYKRVFTTDDKELPLSYLFEVGVYDISEDLPUNVSR-- 342
Oy      495 QAIYSGISLTQLRATQALYYFTDSAECLKKRYHGFQGMNDSTLAVAKIIVE--Y 552

```

```

Db 343 --LQOKKILANISASVKKI--LSEIERLSKDB-KVHHFYE-----PFGVAKBGLY 391
OY 553 AODSNPPOLDLDFIPNNKVKSLGTQMHDEDLVDIIRMEDEKEYIVPT-HNLIRKT 611
Db 392 GDPENKKELELLETFYSKDK--EKLISLKEKENL-----KENOKSIYLLIGENLILKA 444
OY 612 VTGLAGORTKDFHFEIELKNNKOE--LLSOTVKRDKTNLEFROKATINLKHGESLTLQG 669
Db 445 SPLL--EKYAAKGVDVLLSLSEIDAFWPGVNEVDKT--PFKDA-----SHSESLELQG 494
OY 670 LPE-----GSLYLRKETDSEGGKRVNISOEVANATVSRGTGLISDETLAEPENN 716
Db 495 LLELHDEVKQDFKDLMKAFE-ENLKDITKGVLESSLHLSAVALIGDQNNMMAN 547

```

Search completed: June 7, 2001, 00:18:18  
Job time: 225 sec